iScience, Volume *23*

Supplemental Information

Critical Network Structures and Medical Ecology Mechanisms Underlying Human Microbiome-Associated Diseases

Zhanshan (Sam) Ma

Supplemental Information

Supplemental Tables S1-S15 and Figures S1-S3

Table S1. Brief information of the 27 case studies on microbiome associated diseases (MADs). Related to the Transparent Methods and Figs 2-7.

Case No.	Sites	Disease	Treatments (Groups)	References* In "Transparent Methods" section
1	Gut	IBD (Inflammatory Bowel Disease)	Crohn's disease (CD) (18), Ulcerative colitis (UC) (38), Healthy (18)	Papa et al (2012)
2#@			CD (251), UC (324), Healthy (62)	Halfvarson et al (2017)
3		Obesity	Lean (61), Obese (196) and Overweight (24)	Turnbaugh et al (2009)
4		CRC	Cancer (46) vs. Healthy (56)	Wang et al (2012)
5#		HIV	Non-treatment (7) vs. ART treatment (11)	Neff et al (2018)
6			WithART(14), WithoutART(12), Healthy (22)	Lozupone et al (2013)
7			HIV Negative (20) *vs.* Positive (40)	McHardy et al (2013)
8#		Type 1 diabetes and Obesity	Normal T1D (33) vs Normal Healthy (33)	Kim, Jane: https://clinicaltrials.gov/ct2/show/NCT02938806
			Obesity T1D (24) vs. Obesity Healthy (26)	
9		Gout	Disease (41) vs. Healthy (42)	Guo et al (2015)
10		MHE	MHE (25), Cirrhotic (25), Healthy (25)	Zhang et al (2013)
11		Parkinson's Disease	Disease (205) vs. Healthy (133)	Hill-Burns et al 2017)
12		Schizophrenia	Disease (25). vs. Healthy (25)	Dilip Jeste (UC San Diego):https://profiles.ucsd.edu/dilip.jeste
13		Autism, Neurotypical	Autism (88), Neurotypical (41), Healthy (14)	Kang et al (2017)
14		Atherosclerosis	Disease (15) vs. Healthy (15)	Koren et al (2010)
15	Oral	Atherosclerosis	Disease (14) vs. Healthy (15)	Koren et al (2010)
16		Periodontitis	PB (22), PnB (22), Healthy (17)	Abusleme et al (2013)
17			Disease (29), Control (29), Healthy (29)	Griffen et al (2012)
18		Smoking	Smoking (6) vs. Non-smoking (9)	Lazarevic et al (2010)
19			Smoking (74) vs. Non-smoking (72)	Charlson et al (2010)
20	Nostril	Smoking	Smoking (74) vs. Non-smoking (71)	Charlson et al (2010)
21	Skin	Psoriasis	Disease (77), Control (83), Healthy (76)	Alekseyenko et al (2013)
22	Lung	Cystic Fibrosis (CF)	End of Treatment (23) vs. Exacerbation (23)	Fodor et al (2012)
23			Disease (16) vs. Healthy (10)	Blainey et al 2012)
24		HIV	Disease (82) vs. Healthy (77)	Lozupone et al (2013)
25	Vaginal	Bacterial Vaginosis (BV)	ABV, SBV, Healthy	Srinivasan et al. (2012)
26	Semen	Infertile	Abnormal (33), Subnormal (28), Normal (35). Genus level and species level	Weng et al 2014
27@	Milk	Mastitis	Mastitis (4) *vs*. Healthy (16)	Urbaniak (2015)

*These are the same datasets used in Ma, Li & Gotelli (2019) Diversity-disease relationships and shared species analyses for human microbiome-associated diseases. The ISME Journal, https://www.nature.com/articles/s41396-019-0395-y, where a brief description on the datasets can be found.
#: Case No. 2, 5 & 8 datasets were not used for neutral-theoretic analysis because of computational failures.
@: Case No 2 & 27 datasets were not used for core/periphery/skeleton nerwork analysis (computational failures).

*References (Cited in Table S1 for Data Sources)

Papa E, Docktor M, Smillie C, et al. 2012. Non-Invasive Mapping of the Gastrointestinal Microbiota Identifies Children with Inflammatory Bowel Disease. *Plos One*, 7(6):e39242.

Halfvarson J, Brislawn CJ, Lamendella R, et al. 2017. Dynamics of the human gut microbiome in Inflammatory Bowel Disease. *Nature Microbiology*, 2:17004.

Turnbaugh PJ, Hamady M, Yatsunenko T, et al. 2009. A core gut microbiome in obese and lean twins. *Nature*, 457(7228):480.

Wang T, Cai G, Qiu Y, et al. 2012. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *The ISME Journal*, 6(2):320-329.

Neff CP, Krueger O, Xiong K, et al. 2018. Fecal Microbiota Composition Drives Immune Activation in HIV-infected Individuals. *Ebiomedicine*, 30.

Lozupone C, Cota-Gomez A, Palmer B E, et al. 2013. Widespread colonization of the lung by Tropheryma whipplei in HIV infection. *American Journal of Respiratory & Critical Care Medicine*, 187(10):1110-1117.

Mchardy IH, Li X, Tong M, et al. 2013. HIV Infection is associated with compositional and functional shifts in the rectal mucosal microbiota. *Microbiome*, 1(1):26.

Guo Z, Zhang J, Wang Z, et al. 2016. Intestinal Microbiota Distinguish Gout Patients from Healthy Humans. *Sci Rep*, 6:20602.

Zhang Z, Zhai H, Geng J, et al. 2013. Large-scale survey of gut microbiota associated with MHE Via 16S rRNA-based pyrosequencing. *American Journal of Gastroenterology,* 108(10):1601-1611.

Hill-Burns EM, Debelius JW, Morton JT et al. 2017. Parkinson's Disease and PD Medications Have Distinct Signatures of the Gut Microbiome. *Movement Disorders Official Journal of the Movement Disorder Society*, 32(5):739.

Kim, Jane: for Study Case No 8 in Table S1 https://clinicaltrials.gov/ct2/show/NCT02938806

Dilip Jeste (UC San Diego) for Study Case No 12 in Table S1: https://profiles.ucsd.edu/dilip.jeste

Kang DW, Adams JB, Gregory AC, et al. 2017. Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. *Microbiome*, 5(1):10.

Koren O, Klaenhammer TR. 2011. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *PNAS*, 108(Suppl 1):4592-4598.

Koren O, Klaenhammer TR. 2011. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *PNAS*, 108(Suppl 1):4592-4598.

Abusleme L, Dupuy AK, Dutzan N, et al. 2013. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *The ISME Journal*, 7(5):1016-1025.

Griffen AL, Beall CJ, Campbell JH, et al. 2012. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *The ISME Journal*, 6(6):1176.

Lazarevic V, Whiteson K, Hernandez D, et al. 2010. Study of inter- and intra-individual variations in the salivary microbiota. *BMC Genomics*, 11(1):1-11.

Charlson ES, Chen J, Custersallen R, et al. 2010. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *Plos One*, 5(12):e15216.

Charlson ES, Chen J, Custersallen R, et al. 2010. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *Plos One*, 5(12):e15216.

Alekseyenko AV, Perezperez GI, Souza AD, et al. 2013. Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome*,1,1(2013-12-23), 1(1):31-31.

Fodor AA, Klem ER, Gilpin DF, et al. 2012. The Adult Cystic Fibrosis Airway Microbiota Is Stable over Time and Infection Type, and Highly Resilient to Antibiotic Treatment of Exacerbations. *Plos One*, 7(9):e45001.

Blainey PC, Milla CE, Cornfield DN, et al. 2012. Quantitative analysis of the human airway microbial ecology reveals a pervasive signature for cystic fibrosis. *Science Translational Medicine*, 4(153):153ra130.

Lozupone C, Cota-Gomez A, Palmer B E, et al. 2013. Widespread colonization of the lung by Tropheryma whipplei in HIV infection. *American Journal of Respiratory & Critical Care Medicine*, 187(10):1110-1117.

Srinivasan S, Hoffman NG, Morgan MT, et al. 2012. Bacterial Communities in Women with Bacterial Vaginosis: High Resolution Phylogenetic Analyses Reveal Relationships of Microbiota to Clinical Criteria. *Plos One*, 7(6):e37818.

Weng SL, Chiu CM, Lin FM, et al. 2014. Bacterial Communities in Semen from Men of Infertile Couples: Metagenomic Sequencing Reveals Relationships of Seminal Microbiota to Semen Quality. *Plos One*, 9(10):e110152.

Urbaniak C, Mcmillan A, Angelini M, et al. 2014. Effect of chemotherapy on the microbiota and metabolome of human milk, a case report. *Microbiome*, 2(1):24.

Table S3. The percentages (%) that passed neutrality test for Hubbell's classic neutral theory with the 24 datasets of the human MADs (Related to Fig 4)

Sites	Diseases	Treatments		Raw Community Data	Remove rarest species			
					Singleton	M1	M2	M3
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	22.22	11.11	77.78	83.33	66.67
			CD	27.78	5.56	83.33	77.78	72.22
			UC	65.79	44.74	84.21	81.58	71.05
	Obesity	2	Lean	0.00	0.00	65.57	85.25	83.61
			Overweight	4.17	0.00	83.33	83.33	79.17
			Obesity	2.06	0.52	58.76	77.84	81.35
	Cancer	3	Healthy	0.00	0.00	30.36	80.36	75.00
			Cancer	2.17	0.00	45.65	84.78	80.43
	HIV	4	Negative	0.00	0.00	90.48	100.00	100.00
			Treatment	21.43	21.43	100.00	100.00	100.00
			Non-treat	0.00	0.00	83.33	100.00	100.00
		5	Negative	NA	95.00	100.00	100.00	100.00
			Positive	100.00	100.00	100.00	100.00	100.00
	Gout	6	Healthy	38.10	26.19	97.62	100.00	97.62
			Gout	60.98	58.54	100.00	100.00	97.56
	MHE	7	Healthy	12.50	4.17	62.50	65.22	65.22
			Control	16.00	8.00	68.00	80.00	84.00
			MHE	4.17	8.33	70.83	83.33	79.17
	Parkinson's Disease	8	Healthy	89.47	33.08	97.74	96.24	93.98
			PD	88.29	26.47	98.53	95.07	95.57
	Schizophrenia	9	Healthy	12.00	8.00	92.00	100.00	100.00
			Disease	12.00	20.00	100.00	100.00	100.00
	Autism	10	Healthy	92.86	50.00	85.71	85.71	78.57
			Autism	95.45	53.41	94.32	95.45	92.05
			Neurotypical	97.56	73.17	100.00	95.12	95.12
	Atherosclerosis	11	Healthy	0.00	0.00	0.00	46.67	53.33
			Disease	0.00	0.00	6.67	53.33	60.00
Oral	Atherosclerosis	12	Healthy	73.33	26.67	86.67	93.33	93.33
			Disease	42.86	14.29	100.00	100.00	100.00
	Periodontitis	13	Healthy	82.35	52.94	94.12	82.35	82.35
			PB	63.64	18.18	77.27	63.64	68.18
			PnB	40.91	18.18	77.27	77.27	59.09
		14	Healthy	3.45	0.00	17.24	13.79	20.69
			Control	10.34	0.00	20.69	10.34	13.79
			Disease	3.45	0.00	6.90	0.00	6.90
	Smoking	15	Non-Smoker	0.00	0.00	55.56	88.89	77.78
			Smoker	0.00	0.00	100.00	100.00	100.00
		16	Non-Smoker	26.76	38.03	92.96	97.18	94.37
			Smoker	6.85	15.07	87.67	91.78	86.30
		17	Non-Smoker	22.22	0.00	83.33	76.39	72.22
			Smoker	19.18	4.17	69.44	68.06	63.89
Skin	Psoriasis	18	Control	7.23	2.41	32.53	73.49	85.54
			Normal	3.95	1.32	32.89	63.16	73.68

			Lesion	6.49	1.30	40.26	71.43	81.82
Lung	Cystic Fibrosis	19	Treated	100.00	95.45	100.00	95.45	95.24
			Exacerbation	100.00	96.00	100.00	95.65	100.00
		20	Healthy	44.44	11.11	66.67	66.67	55.56
			Disease	56.25	31.25	68.75	68.75	37.50
	HIV	21	Negative	68.42	44.26	78.69	75.00	68.33
			Positive	75.56	53.33	88.89	88.89	84.44
Vaginal	BV	22	Healthy	99.03	95.60	95.60	92.21	93.07
			BV	76.92	47.41	67.24	56.52	49.57
Semen	Infertile (Species level)	23	Normal	8.57	2.86	31.43	62.86	74.29
			Subnormal	0.00	0.00	35.71	75.00	89.29
			Abnormal	0.00	0.00	30.30	66.67	72.73
Milk	Mastitis	24	Healthy	93.75	100.00	100.00	100.00	100.00
			Mastitis	100.00	100.00	100.00	75.00	75.00
Passing Percentages (%)			Health	35.65	26.16	68.75	78.29	77.71
			Diseased	39.13	27.05	75.29	80.88	79.28

Table S4. The *p*-value of Fisher's exact test for the numbers of samples passing the neutrality test between the healthy and diseased datasets (Related to Figs 2-3)

Microbiome	Disease	Index	Healthy	Diseased	Raw Community Data	Remove rarest species			
						Singleton	M1	M2	M3
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	1.000	1.000	1.000	1.000	1.000
			Healthy	UC	0.075	0.078	1.000	1.000	1.000
	Obesity	2	Lean	Overweight	0.291	1.000	0.583	1.000	1.000
			Lean	Obesity	0.576	1.000	0.638	0.744	0.913
	Cancer	3	Healthy	Cancer	0.456	1.000	0.343	0.883	0.881
	HIV	4	Negative	Treatment	0.081	0.081	1.000	1.000	1.000
			Negative	Non-treatment	1.000	1.000	1.000	1.000	1.000
		5	Negative	Positive	NA	1.000	1.000	1.000	1.000
	Gout	6	Healthy	Gout	0.255	0.069	1.000	1.000	1.000
	MHE	7	Healthy	Control	1.000	1.000	1.000	0.663	0.660
			Healthy	MHE	0.611	1.000	0.823	0.657	0.822
	Parkinson's Disease	8	Healthy	PD	0.936	0.352	1.000	1.000	0.936
	Schizophrenia	9	Healthy	Schizophrenia	1.000	0.427	0.843	1.000	1.000
	Autism	10	Healthy	Autism	1.000	1.000	0.837	0.836	0.831
			Healthy	Neurotypical	1.000	0.614	0.823	1.000	0.819
	Atherosclerosis	11	Healthy	Atheroscierosis	1.000	1.000	1.000	1.000	1.000
Oral	Atherosclerosis	12	Healthy	Atheroscierosis	0.540	0.666	1.000	1.000	1.000
	Periodontitis	13	Healthy	PB	0.628	0.199	0.813	0.628	0.807
			Healthy	PnB	0.293	0.199	0.813	1.000	0.618
		14	Healthy	Control	0.613	1.000	1.000	1.000	0.735
			Healthy	Disease	1.000	1.000	0.430	0.116	0.265
	Smoking	15	Non-smoking	Smoking	1.000	1.000	0.692	1.000	1.000
		16	Non-smoking	Smoking	0.008	0.018	0.904	0.905	0.718
		17	Non-smoking	Smoking	0.841	0.245	0.527	0.700	0.695
Skin	Psoriasis	18	Control	Lesion	1.000	1.000	0.541	1.000	0.907
			Normal	Lesion	0.720	1.000	0.536	0.702	0.714
Lung	Cystic Fibrosis (CF)	19	End of Treatment	Exacerbation	1.000	1.000	1.000	1.000	1.000
		20	Healthy	CF	1.000	0.634	1.000	1.000	0.716
	HIV	21	Negative	Positive	0.836	0.610	0.771	0.660	0.550
Vaginal	BV	22	Healthy	BV	0.236	0.002	0.096	0.033	0.004
Semen	Infertile (Species level)	23	Normal	Subnormal	0.256	1.000	0.806	0.695	0.707
			Normal	Abnormal	0.243	1.000	1.000	1.000	1.000
Milk	Mastitis	24	Healthy	Mastitis	1.000	1.000	1.000	1.000	1.000

Table S5. The *p*-values of the permutation tests for the differences in the neutral theory parameters between the healthy and diseased datasets (Related to Figs 2-3)

Micro-biome	Disease	Index	Healthy	Diseased	Raw Community Data		Remove rarest species							
							Singleton		M1		M2		M3	
					θ	m	θ	m	θ	m	θ	m	θ	m
	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	0.542	0.839	0.563	0.239	0.628	0.888	0.815	0.501	0.791	0.406
			Healthy	UC	0.001	0.308	0.000	0.938	0.000	0.993	0.001	0.572	0.002	0.658
	Obesity	2	Lean	Overweight	0.722	0.926	0.707	0.322	0.657	0.030	0.693	0.671	0.650	0.957
			Lean	Obesity	0.081	0.123	0.246	0.408	0.313	0.048	0.225	0.080	0.179	0.493
	Cancer	3	Healthy	Cancer	0.000	0.062	0.043	0.359	0.042	0.827	0.076	0.699	0.321	0.508
	HIV	4	Negative	Treatment	0.052	1.000	0.096	0.325	0.127	0.414	0.175	0.702	0.263	0.606
			Negative	Non-treatment	0.811	0.782	0.471	0.427	0.385	0.699	0.326	0.242	0.274	0.044
		5	Negative	Positive	NA	NA	0.694	0.650	0.460	0.994	0.408	0.866	0.735	0.506
	Gout	6	Healthy	Gout	0.001	0.641	0.001	0.680	0.001	0.401	0.002	0.106	0.003	0.381
	MHE	7	Healthy	Control	0.073	0.670	0.278	0.154	0.132	0.356	0.226	0.653	0.335	0.057
			Healthy	MHE	0.301	0.153	0.255	0.273	0.147	0.959	0.025	0.124	0.028	0.001
	Parkinson's Disease	8	Healthy	PD	0.084	0.149	0.118	0.736	0.077	0.234	0.118	0.471	0.135	0.718
	Schizophrenia	9	Healthy	Schizophrenia	0.307	0.848	0.464	0.832	0.453	0.095	0.397	0.715	0.407	0.788
	Autism	10	Healthy	Autism	0.314	0.487	0.283	0.163	0.300	0.305	0.343	0.842	0.481	0.789
			Healthy	Neurotypical	0.626	0.667	0.695	0.886	0.738	0.626	0.826	0.547	0.782	0.977
	Atherosclerosis	11	Healthy	Atherosclerosis	0.595	0.267	0.775	0.137	0.775	0.217	0.838	0.217	0.870	0.461
Oral	Atherosclerosis	12	Healthy	Atherosclerosis	0.033	0.780	0.026	0.290	0.023	0.780	0.041	0.158	0.026	0.093
	Periodontitis	13	Healthy	PB	0.000	0.604	0.000	0.006	0.000	0.027	0.000	0.333	0.000	0.039
			Healthy	PnB	0.002	0.077	0.000	0.012	0.000	0.408	0.000	0.110	0.001	0.029
		14	Healthy	Control	0.013	0.362	0.012	0.210	0.014	0.265	0.013	0.307	0.008	0.330
			Healthy	Disease	0.000	0.163	0.000	0.163	0.000	0.215	0.000	0.345	0.000	0.353
	Smoking	15	Non-smoking	Smoking	0.607	0.456	0.328	0.088	0.388	0.005	0.689	0.776	0.864	0.272
		16	Non-smoking	Smoking	0.001	0.007	0.055	0.092	0.073	0.842	0.153	0.297	0.057	0.012
		17	Non-smoking	Smoking	0.138	0.100	0.091	0.511	0.069	0.337	0.442	0.790	0.534	0.285
Skin	Psoriasis	18	Control	Lesion	0.889	0.017	0.774	0.040	0.962	0.038	0.951	0.004	0.967	0.348
			Normal	Lesion	0.548	0.937	0.616	0.676	0.789	0.949	0.578	0.811	0.432	0.932
Lung	Cystic Fibrosis (CF)	19	End of Treatment	Exacerbation	0.074	0.195	0.492	1.000	0.453	0.992	0.227	0.411	0.363	0.442
		20	Healthy	CF	0.000	0.388	0.000	0.207	0.000	0.251	0.049	0.032	0.049	0.037
	HIV	21	Negative	Positive	0.286	0.422	0.697	0.486	0.645	0.730	0.828	0.571	0.843	0.571
Vaginal	BV	22	Healthy	BV	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Semen	Infertile (Species level)	23	Normal	Subnormal	0.550	0.437	0.429	0.645	0.470	0.645	0.462	0.174	0.268	0.726
			Normal	Abnormal	0.705	0.922	0.592	0.542	0.617	0.212	0.464	0.271	0.335	0.145
Milk	Mastitis	24	Healthy	Mastitis	0.494	0.554	1.000	0.800	0.800	0.533	1.000	0.800	0.374	0.304
Significant difference rate (%)					33.33	9.09	30.30	12.12	30.30	18.18	30.30	9.09	30.30	21.21

Table S7. The shared core/periphery analysis between the healthy and diseased treatments based on the 1000 sets of permutated core/periphery networks (CPN) built with FDR control and p=0.001 (Related to Figs 6-7)

Microbiome	Disease	Index	Healthy	Diseased	Core			Periphery		
					Obs.	Perm.	*p-value*	Obs.	Perm.	*p-value*
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	1	1.618	0.545	7	9.674	0.282
			Healthy	UC	1	0.900	0.776	9	6.196	0.898
	Obesity	3	Lean	Overweight	1	1.895	0.505	12	7.495	0.932
			Lean	Obesity	10	9.781	0.597	23	20.753	0.720
	Cancer	4	Healthy	Cancer	9	11.796	0.249	11	15.48	0.160
	HIV	5	Negative	Treatment	22	28.498	0.178	43	37.691	0.792
			Negative	Non-treatment	24	25.658	0.456	27	30.546	0.357
		6	Negative	Treatment	15	26.606	0.029	25	33.909	0.107
			Negative	Non-treatment	11	26.427	0.002	23	32.571	0.106
		7	Negative	Positive	22	13.173	0.957	18	18.304	0.541
	T1D (Lean)	8	Healthy	T1D	10	8.496	0.763	19	16.18	0.792
	T1D (Obesity)		Healthy	T1D	20	15.005	0.902	28	24.807	0.760
	Gout	9	Healthy	Gout	1	5.43	0.035	17	17.809	0.492
	MHE	10	Healthy	MHE	6	4.86	0.772	6	13.842	0.026
			Control	MHE	3	4.702	0.342	12	13.68	0.427
	Parkinson's Disease	11	Healthy	PD	38	45.7	0.170	40	47.696	0.177
	Schizophrenia	12	Healthy	Schizophrenia	30	29.815	0.547	26	34.969	0.077
	Autism	13	Healthy	Autism	3	1.97	0.865	9	6.978	0.768
			Healthy	Neurotypical	0	0.706	0.679	4	4.875	0.580
	Atherosclerosis	14	Healthy	Atherosclerosis	29	27.573	0.671	23	30.778	0.102
Oral	Atherosclerosis	15	Healthy	Atherosclerosis	0	0.523	0.589	2	4.616	0.170
	Periodontitis	16	Healthy	PB	1	0.592	0.865	2	4.145	0.257
			Healthy	PnB	0	0.619	0.605	2	3.978	0.266
		17	Healthy	Disease	15	6.815	0.928	14	6.387	0.939
			Control	Disease	0	6.799	0.092	0	6.244	0.085
	Smoking	18	Non-smoking	Smoking	4	2.368	0.887	13	14.387	0.438
		19	Non-smoking	Smoking	20	34.991	0.019	20	34.701	0.008
		20	Non-smoking	Smoking	0	0.957	0.522	16	11.421	0.905
Skin	Psoriasis	21	Control	Lesion	84	112.266	0.014	94	116.43	0.032
			Normal	Lesion	109	105.23	0.644	106	110.03	0.393
Lung	Cystic Fibrosis (CF)	22	End of Treatment	Exacerbation	0	8.69	0.011	0	27.068	0.000
		23	Healthy	CF	0	0.549	0.644	1	2.016	0.447
	HIV	24	Negative	Positive	181	168.219	0.605	118	105.99	0.741
Vaginal	BV	25	Healthy	BV	0	13.183	0.003	0	23.395	0.000
Semen	Infertile (Genus level)	26	Normal	Subnormal	1	59.733	0.000	0	57.235	0.000
			Normal	Abnormal	0	63.879	0.000	4	58.901	0.000
	Infertile Species level)		Normal	Subnormal	69	64.506	0.658	63	67.569	0.378
			Normal	Abnormal	67	65.134	0.584	51	69.039	0.066
Percentage of Comparisons with Reduced Shared Skeletons					23.68% (9/38)			18.42% (7/38)		
Percentage of Comparisons without the change					76.32% (29/38)			81.58% (31/38)		

*Obs.=The Observed value. Perm.=The mean value from 1000 permutated networks. p-value is from permutation test.

Table S8. The *p*-values from the permutation test for the differences in the core/periphery network (CPN) properties between the H (healthy) & D (diseased) treatments with 1000 sets of permutated CPNs (*p*-value<0.05, indicating significant difference) (Related to Figs 6-7)

Microbiome	Disease	Index	Healthy	Diseased	ρ	C/ (C+P)	B11	B12 (21)	B22	Nestedness (S)
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	0.00	0.00	0.00	1.00	0.00	1.00
			Healthy	UC	0.00	1.00	1.00	0.00	0.00	0.00
	Obesity	3	Lean	Overweight	0.00	0.00	0.00	1.00	1.00	0.00
			Lean	Obesity	0.00	0.00	0.00	1.00	0.00	0.00
	Cancer	4	Healthy	Cancer	0.00	0.00	0.00	0.00	0.00	0.00
	HIV	5	Negative	Treatment	0.00	0.00	0.00	0.00	1.00	0.00
			Negative	Non-treatment	1.00	1.00	1.00	0.00	0.00	0.00
		6	Negative	Treatment	0.00	1.00	0.00	0.00	0.00	0.00
			Negative	Non-treatment	1.00	0.00	1.00	1.00	0.00	1.00
		7	Negative	Positive	0.00	1.00	0.00	0.00	1.00	0.00
	T1D (Lean)	8	Healthy	T1D	0.00	0.00	0.00	0.00	0.00	0.00
	T1D (Obesity)		Healthy	T1D	1.00	1.00	0.00	0.00	0.00	0.00
	Gout	9	Healthy	Gout	1.00	1.00	1.00	0.00	0.00	1.00
	MHE	10	Healthy	MHE	1.00	1.00	1.00	0.00	0.00	1.00
			Control	MHE	1.00	1.00	1.00	0.00	1.00	1.00
	Parkinson's Disease	11	Healthy	PD	0.00	0.00	0.00	0.00	0.00	0.00
	Schizophrenia	12	Healthy	Schizophrenia	0.00	0.00	1.00	1.00	1.00	1.00
	Autism	13	Healthy	Autism	1.00	1.00	1.00	1.00	0.00	1.00
			Healthy	Neurotypical	1.00	1.00	1.00	1.00	0.00	1.00
	Atherosclerosis	14	Healthy	Atherosclerosis	1.00	1.00	1.00	1.00	0.00	1.00
Oral	Atherosclerosis	15	Healthy	Atherosclerosis	0.00	0.00	0.00	1.00	0.00	0.00
	Periodontitis	16	Healthy	PB	1.00	1.00	0.00	1.00	1.00	1.00
			Healthy	PnB	0.00	0.00	0.00	1.00	1.00	0.00
		17	Healthy	Disease	0.00	1.00	0.00	0.00	0.00	0.00
			Control	Disease	0.00	1.00	1.00	1.00	0.00	1.00
	Smoking	18	Non-smoking	Smoking	1.00	1.00	1.00	1.00	1.00	1.00
		19	Non-smoking	Smoking	0.00	1.00	0.00	0.00	0.00	0.00
		20	Non-smoking	Smoking	1.00	0.00	1.00	0.00	1.00	1.00
Skin	Psoriasis	21	Control	Lesion	0.00	1.00	1.00	1.00	1.00	0.00
			Normal	Lesion	0.00	1.00	1.00	0.00	0.00	0.00
Lung	Cystic Fibrosis (CF)	22	End of Treatment	Exacerbation	0.00	0.00	1.00	1.00	0.00	0.00
		23	Healthy	CF	1.00	1.00	1.00	1.00	0.00	0.00
	HIV	24	Negative	Positive	0.00	1.00	0.00	0.00	0.00	0.00
Vaginal	BV	25	Healthy	BV	0.00	0.00	0.00	1.00	0.00	0.00
Semen	Infertile (Genus level)	26	Normal	Subnormal	0.00	1.00	0.00	1.00	0.00	0.00
			Normal	Abnormal	0.00	0.00	1.00	1.00	0.00	0.00
	Infertile Species level)		Normal	Subnormal	0.00	1.00	0.00	0.00	0.00	0.00
			Normal	Abnormal	1.00	1.00	1.00	1.00	1.00	1.00
Percentage of Significant Differences (%)					59.18	44.90	51.02	46.94	63.27	57.14

Here the algorithm for permutation test is as follows:
OP=|Observed Property Difference between H & D|
PP=1000 values of |H-D| from permutated networks
N=number of permutations satisfying PP≤OP
Pseudo-P-value=N/1000

Table S9. The shared skeleton analysis between the healthy and diseased treatments based on the 1000 sets of permutated high-salience skeleton networks (HSNs) built with FDR control and p=0.001 (HSS>0) (Related to Figs 6-7)

Microbiome	Disease	Index	Healthy	Diseased	HSS (High Salience Skeleton) Value>0		
					Obs.	**P**erm.	***p**-value*
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	49	143.27	0.111
			Healthy	UC	80	55.44	0.785
	Obesity	3	Lean	Overweight	166	111.85	0.803
			Lean	Obesity	1132	1072.96	0.605
	Cancer	4	Healthy	Cancer	630	1185.00	0.026
	HIV	5	Negative	Treatment	NA	NA	NA
			Negative	Non-treatment	NA	NA	NA
		6	Negative	Treatment	2718	6057.995	0.007
			Negative	Non-treatment	1934	5855.577	0.003
		7	Negative	Positive	2049	1495.90	0.839
	T1D (Lean)	8	Healthy	T1D	1059	923.64	0.720
	T1D (Obesity)		Healthy	T1D	3393	2411.57	0.965
	Gout	9	Healthy	Gout	407	748.53	0.046
	MHE	10	Healthy	MHE	172	489.41	0.041
			Control	MHE	309	478.13	0.206
	Parkinson's Disease	11	Healthy	PD	3630	6724.11	0.015
	Schizophrenia	12	Healthy	Schizophrenia	4861	6664.09	0.067
	Autism	13	Healthy	Autism	43	89.80	0.425
			Healthy	Neurotypical	24	46.23	0.662
	Atherosclerosis	14	Healthy	Atherosclerosis	4288	5592.62	0.083
Oral	Atherosclerosis	15	Healthy	Atherosclerosis	2	26.69	0.031
	Periodontitis	16	Healthy	PB	1	23.57	0.057
			Healthy	PnB	1	23.11	0.047
		17	Healthy	Disease	1353	441.901	0.944
			Control	Disease	0	437.733	0.043
	Smoking	18	Non-smoking	Smoking	279	353.64	0.388
		19	Non-smoking	Smoking	1707	4029.50	0.029
		20	Non-smoking	Smoking	153	147.17	0.577
Skin	Psoriasis	21	Control	Lesion	37098	62519.264	0.002
			Normal	Lesion	64145	55303.513	0.833
Lung	Cystic Fibrosis (CF)	22	End of Treatment	Exacerbation	0	456.75	0.000
		23	Healthy	CF	0	6.78	0.202
	HIV	24	Negative	Positive	19746	33750.779	0.346
Vaginal	BV	25	Healthy	BV	0	23.359	0.000
Semen	Infertile (Genus level)	26	Normal	Subnormal	3	654.33	0.000
			Normal	Abnormal	6	561.187	0.000
	Infertile Species level)		Normal	Subnormal	26291	29221.901	0.402
			Normal	Abnormal	24225	29292.624	0.249
Percentage of Comparisons with Reduced Shared Skeletons					39.47%(15/38)		
Percentage of Comparisons without the change					60.53%(23/38)		

*In general, the observed shared skeletons (*OSS*) should be smaller or, at most, equal with the permutated shared skeletons (*PSS*), *i.e.*, $OSS \leq PSS$. In other words, the case of $O>P$ is unlikely and corresponding p-value would be p=1.

Table S10. The *p*-values from the permutation test for the differences in the high salience skeleton network (HSN) properties between the H & D treatments with 1000 sets of permutated HSNs (*p* value<0.05, indicating significant difference) (Related to Figs 6-7)

Microbiome	Disease	Index	Healthy	Diseased	Links (%)	Max	Skewness	Kurtosis	r_{HSS}
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	1.00	0.00	1.00	0.00	0.00
			Healthy	UC	1.00	0.00	1.00	1.00	0.00
	Obesity	3	Lean	Overweight	1.00	1.00	1.00	0.00	0.00
			Lean	Obesity	0.00	0.00	0.00	0.00	0.00
	Cancer	4	Healthy	Cancer	0.00	0.00	1.00	1.00	0.00
	HIV	5	Negative	Treatment	0.00	0.00	1.00	1.00	0.00
			Negative	Non-treatment	0.00	0.00	1.00	0.00	0.00
		6	Negative	Treatment	0.00	0.00	0.00	0.00	0.00
			Negative	Non-treatment	0.00	1.00	0.00	1.00	0.00
		7	Negative	Positive	1.00	1.00	1.00	0.00	0.00
	T1D (Lean)	8	Healthy	T1D	1.00	1.00	1.00	0.00	1.00
	T1D (Obesity)		Healthy	T1D	0.00	0.00	1.00	1.00	1.00
	Gout	9	Healthy	Gout	0.00	0.00	1.00	1.00	1.00
	MHE	10	Healthy	MHE	1.00	0.00	1.00	1.00	0.00
			Control	MHE	0.00	1.00	0.00	0.00	0.00
	Parkinson's Disease	11	Healthy	PD	0.00	0.00	1.00	1.00	0.00
	Schizophrenia	12	Healthy	Schizophrenia	0.00	0.00	0.00	0.00	0.00
	Autism	13	Healthy	Autism	0.00	0.00	0.00	0.00	0.00
			Healthy	Neuro-typical	1.00	1.00	0.00	0.00	1.00
	Atherosclerosis	14	Healthy	Atherosclerosis	0.00	0.00	0.00	0.00	0.00
Oral	Atherosclerosis	15	Healthy	Atherosclerosis	1.00	0.00	1.00	1.00	0.00
	Periodontitis	16	Healthy	PB	0.00	0.00	0.00	0.00	0.00
			Healthy	PnB	0.00	0.00	1.00	1.00	0.00
		17	Healthy	Disease	1.00	1.00	0.00	0.00	0.00
			Control	Disease	1.00	1.00	0.00	0.00	0.00
	Smoking	18	Non-smoking	Smoking	1.00	1.00	1.00	0.00	1.00
		19	Non-smoking	Smoking	0.00	0.00	0.00	0.00	1.00
		20	Non-smoking	Smoking	1.00	1.00	1.00	1.00	1.00
Skin	Psoriasis	21	Control	Lesion	0.00	1.00	1.00	1.00	0.00
			Normal	Lesion	1.00	1.00	0.00	0.00	1.00
Lung	Cystic Fibrosis (CF)	22	End of Treatment	Exacerbation	1.00	0.00	1.00	1.00	0.00
		23	Healthy	CF	0.00	0.00	0.00	0.00	1.00
	HIV	24	Negative	Positive	0.00	0.00	0.00	0.00	0.00
Vaginal	BV	25	Healthy	BV	0.00	0.00	0.00	0.00	0.00
Semen	Infertile (Genus level)	26	Normal	Subnormal	1.00	0.00	0.00	0.00	0.00
			Normal	Abnormal	1.00	0.00	0.00	0.00	0.00
	Infertile Species level)		Normal	Subnormal	1.00	1.00	1.00	1.00	1.00
			Normal	Abnormal	1.00	0.00	1.00	0.00	0.00
% With significant differences					52.63	65.79	47.37	63.16	73.68

Table S11. The *observed* core/periphery parameters in the core-periphery networks for the 25 MAD (microbiome associated disease) datasets with FDR control and *p*-value=0.001 (Related to Figs 6-7)

Sites	Diseases	Treatments		*r*	C/(C+P)	Density Matrix			Nestedness (S)
						B11	B12 (21)	B22	
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	0.179	0.310	0.098	0.001	0.022	0.020
			CD	0.254	0.174	0.210	0.000	0.016	0.018
			UC	0.246	0.344	0.119	0.003	0.014	0.021
	Obesity	3	Lean	0.405	0.250	0.277	0.004	0.013	0.026
			Overweight	0.236	0.283	0.129	0.007	0.012	0.019
			Obesity	0.042	0.478	0.012	0.004	0.005	0.009
	Cancer	4	Healthy	0.027	0.470	0.006	0.002	0.003	0.003
			Cancer	0.079	0.406	0.021	0.003	0.004	0.006
	HIV	5	Negative	0.074	0.459	0.019	0.003	0.005	0.007
			Treatment	0.104	0.498	0.031	0.006	0.004	0.012
			Non-treatment	0.086	0.465	0.028	0.007	0.006	0.011
		6	Negative	0.117	0.446	0.039	0.007	0.006	0.013
			Treatment	0.082	0.442	0.026	0.005	0.007	0.010
			Non-treatment	0.091	0.413	0.034	0.007	0.008	0.012
		7	Negative	0.180	0.444	0.066	0.006	0.007	0.018
			Positive	0.064	0.434	0.017	0.003	0.006	0.007
	T1D (Lean)	8	Normal-Healthy	0.098	0.426	0.029	0.003	0.007	0.009
			Normal-T1D	0.068	0.480	0.018	0.003	0.007	0.007
	T1D (Obesity)		Obesity-Healthy	0.068	0.465	0.016	0.002	0.006	0.006
			Obesity-T1D	0.068	0.469	0.020	0.004	0.008	0.009
	Gout	9	Healthy	0.178	0.411	0.071	0.006	0.010	0.018
			Gout	0.118	0.436	0.041	0.004	0.011	0.013
	MHE	10	Healthy	0.174	0.310	0.081	0.005	0.011	0.015
			Control	0.182	0.345	0.087	0.006	0.014	0.019
			MHE	0.225	0.287	0.120	0.003	0.014	0.018
	Parkinson's Disease	11	Healthy	0.083	0.429	0.023	0.004	0.005	0.013
			PD	0.037	0.507	0.008	0.003	0.003	0.007
	Schizophrenia	12	Healthy	0.058	0.490	0.017	0.005	0.005	0.008
			Disease	0.059	0.460	0.018	0.005	0.006	0.008
	Autism	13	Healthy	0.301	0.331	0.218	0.013	0.030	0.043
			Autism	0.176	0.345	0.075	0.005	0.010	0.023
			Neurotypical	0.187	0.374	0.086	0.004	0.017	0.021
	Atherosclerosis	14	Healthy	0.054	0.483	0.015	0.004	0.006	0.007
			Disease	0.039	0.500	0.012	0.004	0.006	0.006
Oral	Atherosclerosis	15	Healthy	0.430	0.218	0.470	0.000	0.043	0.047
			Disease	0.164	0.340	0.076	0.000	0.019	0.017
	Periodontitis	16	Healthy	0.336	0.255	0.286	0.000	0.035	0.037
			PB	0.186	0.289	0.128	0.000	0.032	0.027
			PnB	0.612	0.136	1.000	0.000	0.034	0.042
		17	Healthy	0.113	0.496	0.153	0.068	0.093	0.095
			Control	0.027	0.518	0.032	0.023	0.021	0.025
			Disease	0.013	0.513	0.021	0.024	0.029	0.024
	Smoking	18	Non-Smoker	0.187	0.378	0.114	0.016	0.023	0.032
			Smoker	0.296	0.383	0.205	0.017	0.031	0.049
		19	Non-Smoker	0.170	0.391	0.082	0.011	0.012	0.037
			Smoker	0.094	0.437	0.032	0.006	0.008	0.019
		20	Non-Smoker	0.382	0.206	0.324	0.001	0.020	0.031
			Smoker	0.491	0.188	0.431	0.006	0.013	0.030
Skin	Psoriasis	21	Control	0.036	0.497	0.010	0.004	0.003	0.007
			Normal	0.032	0.491	0.008	0.003	0.003	0.006

			Lesion	0.028	0.501	0.008	0.004	0.003	0.007
Lung	Cystic Fibrosis	22	Treated	0.000	1.000	1.000	0.000	0.000	1.000
			Exacerbation	0.471	0.250	1.000	0.000	0.200	0.143
		23	Healthy	0.681	0.207	1.000	0.000	0.063	0.076
			Disease	0.608	0.167	1.000	0.000	0.053	0.061
	HIV	24	Negative	0.089	0.496	0.050	0.020	0.012	0.045
			Positive	0.188	0.495	0.146	0.045	0.024	0.090
Vaginal	BV	25	Healthy	0.387	0.353	0.400	0.000	0.109	0.092
			BV	0.632	0.500	1.000	0.000	1.000	0.333
Semen	Infertile (Genus level)	26	Normal	0.326	0.250	0.400	0.000	0.076	0.066
			Subnormal	0.625	0.227	1.000	0.000	0.103	0.104
			Abnormal	0.348	0.214	0.400	0.000	0.048	0.045
	Infertile (Species level)		Normal	0.051	0.501	0.014	0.005	0.004	0.007
			Subnormal	0.082	0.461	0.027	0.007	0.005	0.011
			Abnormal	0.049	0.488	0.016	0.007	0.005	0.008
Healthy		**Mean**		0.175	0.419	0.178	0.010	0.023	0.051
		Standard Error		0.026	0.020	0.042	0.003	0.004	0.022
Diseased		**Mean**		0.216	0.374	0.238	0.006	0.052	0.037
		Standard Error		0.033	0.020	0.060	0.001	0.028	0.010

Table S12. The *observed* statistical network properties of the high-salience skeleton networks for the 25 MAD (microbiome associated disease) datasets with FDR control and *p*-value=0.001 (Related to Figs 6-7)

Sites	Diseases	Treatments		Statistics of HSS						Assortativity
				Links (%)	Max	Mean	Std. Error	Skewness	Kurtosis	r_{HSS}
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	99.985	0.026	0.017	0.000	-33.349	3334.500	-0.009
			CD	99.945	0.035	0.023	0.001	-17.459	911.999	-0.012
			UC	99.984	0.019	0.013	0.000	-31.949	3059.750	-0.006
	Obesity	3	Lean	99.948	0.026	0.013	0.000	-7.085	973.275	-0.007
			Overweight	99.994	0.016	0.011	0.000	-53.842	8695.000	-0.005
			Obesity	73.168	0.519	0.004	0.020	22.358	519.682	-0.002
	Cancer	4	Healthy	99.992	0.008	0.003	0.000	29.987	8911.597	-0.002
			Cancer	100.000	0.004	0.004	0.000	0.000	0.000	-0.002
	HIV	5	Negative	99.959	0.015	0.003	0.000	50.530	5104.754	-0.002
			Treatment	99.998	0.005	0.002	0.000	-72.261	24815.977	-0.001
			Non-treat	99.950	0.017	0.002	0.000	54.377	5780.542	-0.001
		6	Negative	99.995	0.007	0.004	0.000	-48.101	10994.593	-0.002
			Treatment	99.999	0.007	0.004	0.000	-129.184	50063.500	-0.002
			Non-treat	99.999	0.007	0.004	0.000	-129.184	50063.500	-0.002
		7	Negative	99.995	0.010	0.005	0.000	-37.779	9847.053	-0.003
			Positive	99.973	0.015	0.006	0.000	-3.581	2036.490	-0.003
	T1D (Lean)	8	Normal-Healthy	99.998	0.010	0.007	0.000	-87.327	22876.000	-0.003
			Normal-T1D	99.996	0.010	0.007	0.000	-62.362	11664.500	-0.003
	T1D (Obesity)		Obesity-Healthy	100.000	0.006	0.006	0.000	0.000	0.000	-0.003
			Obesity-T1D	99.998	0.008	0.006	0.000	-103.204	31951.000	-0.003
	Gout	9	Healthy	99.949	0.030	0.008	0.000	27.417	3262.266	-0.004
			Gout	99.986	0.017	0.008	0.000	-6.210	3660.833	-0.004
	MHE	10	Healthy	100.000	0.010	0.010	0.000	0.000	0.000	-0.005
			Control	99.996	0.014	0.009	0.000	-63.368	12044.500	-0.005
			MHE	100.000	0.013	0.013	0.000	0.000	0.000	-0.007
	Parkinson's Disease	11	Healthy	88.116	0.332	0.005	0.012	22.326	540.969	-0.003
			PD	59.691	0.638	0.003	0.024	24.250	610.921	-0.001
	Schizophrenia	12	Healthy	100.000	0.004	0.004	0.000	0.000	0.000	-0.002
			Disease	99.999	0.005	0.004	0.000	-114.004	38988.500	-0.002
	Autism	13	Healthy	99.992	0.019	0.013	0.000	-45.183	6122.500	-0.006
			Autism	91.270	0.301	0.009	0.013	18.661	378.342	-0.004
			Neurotypical	99.947	0.040	0.011	0.000	16.341	2387.524	-0.006
	Atherosclerosis	14	Healthy	99.999	0.006	0.004	0.000	-141.020	59657.500	-0.002
			Disease	99.999	0.005	0.004	0.000	-163.247	79947.000	-0.002
Oral	Atherosclerosis	15	Healthy	100.000	0.036	0.036	0.000	0.000	0.000	-0.019
			Disease	100.000	0.019	0.019	0.000	0.000	0.000	-0.010
	Periodontitis	16	Healthy	100.000	0.036	0.036	0.000	0.000	0.000	-0.019
			PB	99.899	0.067	0.044	0.002	-12.865	494.500	-0.023
			PnB	99.894	0.068	0.045	0.002	-12.576	472.500	-0.023
		17	Healthy	99.999	0.008	0.005	0.000	-115.039	39700.000	-0.003
			Control	100.000	0.002	0.002	0.000	-350.019	367538.500	-0.001
			Disease	100.000	0.002	0.001	0.000	-409.198	502326.000	-0.001
	Smoking	18	Non-Smoker	99.995	0.014	0.010	0.000	-60.193	10867.500	-0.005
			Smoker	99.993	0.017	0.011	0.000	-50.379	7612.000	-0.006
		19	Non-Smoker	70.997	0.520	0.008	0.029	14.357	223.799	-0.004
			Smoker	91.622	0.202	0.006	0.008	18.791	387.874	-0.003
		20	Non-Smoker	99.029	0.108	0.020	0.003	7.964	211.591	-0.010
			Smoker	98.322	0.121	0.013	0.004	15.895	348.415	-0.007

Skin	Psoriasis	21	Control	96.188	0.133	0.002	0.003	38.406	1640.442	-0.001
			Normal	94.728	0.170	0.002	0.003	39.778	1789.015	-0.001
			Lesion	94.563	0.149	0.002	0.003	34.358	1341.691	-0.001
Lung	Cystic Fibrosis	22	Treated	100.000	1.000	1.000	NA	NA	NA	-1.000
			Exacerbation	100.000	0.250	0.250	0.000	0.000	0.000	-0.143
		23	Healthy	100.000	0.069	0.069	0.000	0.000	0.000	-0.036
			Disease	99.770	0.100	0.067	0.004	-8.544	217.000	-0.034
	HIV	24	Negative	13.101	0.946	0.001	0.020	29.803	1053.016	-0.001
			Positive	92.261	0.365	0.003	0.004	52.143	4337.595	-0.001
Vaginal	BV	25	Healthy	99.265	0.176	0.118	0.012	-4.814	67.500	-0.062
			BV	100.000	0.500	0.500	0.000	0.000	0.000	-0.333
Semen	Infertile (Genus level)	26	Normal	99.474	0.150	0.100	0.009	-5.672	94.500	-0.053
			Subnormal	100.000	0.091	0.091	0.000	0.000	0.000	-0.048
			Abnormal	100.000	0.071	0.071	0.000	0.000	0.000	-0.037
	Infertile (Species level)		Normal	99.993	0.012	0.002	0.000	105.660	33899.677	-0.001
			Subnormal	99.997	0.007	0.002	0.000	112.384	35971.995	-0.001
			Abnormal	99.998	0.006	0.002	0.000	102.264	55854.225	-0.001
Healthy		**Mean**		95.506	0.126	0.049	0.003	-25.693	22119.667	-0.041
		Std. Err.		2.924	0.045	0.032	0.001	14.768	12064.204	0.032
Diseased		**Mean**		97.065	0.109	0.037	0.003	-22.651	25354.820	-0.022
		Std. Err.		1.429	0.029	0.016	0.001	15.350	14832.102	0.010

Table S13. The *observed* number of shared core nodes between the healthy (H) and diseased (D) treatments in the 25 MAD (microbiome associated disease) datasets (Related to Figs 6-7)

Microbiome	Disease	Index	Healthy	Diseased	Number of Core Nodes in Healthy	Number of Core Nodes in Diseased	Shared Core Nodes	% of Core Nodes in terms of Healthy	% of Core Nodes in terms of Diseased
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	36	15	1	2.78	6.67
			Healthy	UC	36	54	1	2.78	1.85
	Obesity	3	Lean	Overweight	38	53	1	2.63	1.89
			Lean	Obesity	38	257	10	26.32	3.89
	Cancer	4	Healthy	Cancer	280	182	9	3.21	4.95
	HIV	5	Negative	Treatment	278	413	22	7.91	5.33
			Negative	Non-treatment	278	382	24	8.63	6.28
		6	Negative	Treatment	246	198	15	6.10	7.58
			Negative	Non-treatment	246	185	11	4.47	5.95
		7	Negative	Positive	175	145	22	12.57	15.17
	T1D (Lean)	8	Healthy	T1D	129	147	10	7.75	6.80
	T1D (Obesity)		Healthy	T1D	167	168	20	11.98	11.90
	Gout	9	Healthy	Gout	109	105	1	0.92	0.95
	MHE	10	Healthy	MHE	61	43	6	9.84	13.95
			Control	MHE	76	43	3	3.95	6.98
	Parkinson's Disease	11	Healthy	PD	164	386	38	23.17	9.84
	Schizophrenia	12	Healthy	Schizophrenia	271	257	30	11.07	11.67
	Autism	13	Healthy	Autism	52	79	3	5.77	3.80
			Healthy	Neurotypical	52	65	0	0.00	0.00
	Atherosclerosis	14	Healthy	Atherosclerosis	236	283	29	12.29	10.25
Oral	Atherosclerosis	15	Healthy	Atherosclerosis	12	35	0	0.00	0.00
	Periodontitis	16	Healthy	PB	14	13	1	7.14	7.69
			Healthy	PnB	14	6	0	0.00	0.00
		17	Healthy	Disease	198	728	15	7.58	2.06
			Control	Disease	628	728	0	0.00	0.00
	Smoking	18	Non-smoking	Smoking	79	67	4	5.06	5.97
		19	Non-smoking	Smoking	97	156	20	20.62	12.82
		20	Non-smoking	Smoking	21	28	0	0.00	0.00
Skin	Psoriasis	21	Control	Lesion	446	508	84	18.83	16.54
			Normal	Lesion	499	508	109	21.84	21.46
Lung	Cystic Fibrosis (CF)	22	End of Treatment	Exacerbation	2	2	0	0.00	0.00
		23	Healthy	CF	6	5	0	0.00	0.00
	HIV	24	Negative	Positive	804	338	181	22.51	53.55
Vaginal	BV	25	Healthy	BV	6	2	0	0.00	0.00
Semen	Infertile (Genus level)	26	Normal	Subnormal	5	5	1	20.00	20.00
			Normal	Abnormal	5	6	0	0.00	0.00
	Infertile Species level)		Normal	Subnormal	429	473	69	16.08	14.59
			Normal	Abnormal	429	519	67	15.62	12.91
Mean & Standard Error			**Mean**		175.32	199.66	21.24	8.41	7.98
			Standard Error		31.13	33.59	5.98	1.29	1.58

Table S14. The *observed* number of shared periphery nodes between the healthy and diseased treatments in the 25 MAD (microbiome associated disease) datasets (Related to Figs 6-7)

Microbiome	Disease	Index	Healthy	Diseased	Number of Periphery Nodes in Healthy	Number of Periphery Nodes in Diseased	Shared Periphery Nodes	% of Periphery Nodes in terms of Healthy	% of Periphery Nodes in terms of Diseased
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	80	71	7	8.75	9.86
			Healthy	UC	80	103	9	11.25	8.74
	Obesity	3	Lean	Overweight	114	134	12	10.53	8.96
			Lean	Obesity	114	281	23	20.18	8.19
	Cancer	4	Healthy	Cancer	316	266	11	3.48	4.14
	HIV	5	Negative	Treatment	328	416	43	13.11	10.34
			Negative	Non-treatment	328	440	27	8.23	6.14
		6	Negative	Treatment	306	250	25	8.17	10.00
			Negative	Non-treatment	306	263	23	7.52	8.75
		7	Negative	Positive	219	189	18	8.22	9.52
	T1D (Lean)	8	Healthy	T1D	174	159	19	10.92	11.95
	T1D (Obesity)		Healthy	T1D	192	190	28	14.58	14.74
	Gout	9	Healthy	Gout	156	136	17	10.90	12.50
	MHE	10	Healthy	MHE	136	107	6	4.41	5.61
			Control	MHE	144	107	12	8.33	11.21
	Parkinson's Disease	11	Healthy	PD	218	376	40	18.35	10.64
	Schizophrenia	12	Healthy	Schizophrenia	282	302	26	9.22	8.61
	Autism	13	Healthy	Autism	105	150	9	8.57	6.00
			Healthy	Neurotypical	105	109	4	3.81	3.67
	Atherosclerosis	14	Healthy	Atherosclerosis	253	283	23	9.09	8.13
Oral	Atherosclerosis	15	Healthy	Atherosclerosis	43	68	2	4.65	2.94
	Periodontitis	16	Healthy	PB	41	32	2	4.88	6.25
			Healthy	PnB	41	38	2	4.88	5.26
		17	Healthy	Disease	201	690	14	6.97	2.03
			Control	Disease	585	690	0	0.00	0.00
	Smoking	18	Non-smoking	Smoking	130	108	13	10.00	12.04
		19	Non-smoking	Smoking	151	201	20	13.25	9.95
		20	Non-smoking	Smoking	81	121	16	19.75	13.22
Skin	Psoriasis	21	Control	Lesion	452	506	94	20.80	18.58
			Normal	Lesion	517	506	106	20.50	20.95
Lung	Cystic Fibrosis (CF)	22	End of Treatment	Exacerbation	0	6	0	NA	0.00
		23	Healthy	CF	23	25	1	4.35	4.00
	HIV	24	Negative	Positive	816	345	118	14.46	34.20
Vaginal	BV	25	Healthy	BV	11	2	0	0.00	0.00
Semen	Infertile (Genus level)	26	Normal	Subnormal	15	17	0	0.00	0.00
			Normal	Abnormal	15	22	4	26.67	18.18
	Infertile Species level)		Normal	Subnormal	428	553	63	14.72	11.39
			Normal	Abnormal	428	545	51	11.92	9.36
Mean & Standard Error			**Mean**		208.79	231.76	23.37	10.15	9.11
			Standard Error		29.40	31.41	4.65	1.02	1.06

Table S15. The *observed* number of shared skeletons with *HSS*>0 between the healthy and diseased treatments in the 25 MAD (microbiome associated disease) datasets (Related to Figs 6-7)

Microbiome	Disease	Index	Healthy	Diseased	Number of Edges in Healthy	Number of Edges Diseased	Shared Edges	% of Edges in terms of Healthy	% of Edges in terms of Diseased
Gut	IBD (Inflammatory Bowel Disease)	1	Healthy	CD	6669	3653	49	0.73	1.34
			Healthy	UC	6669	12244	80	1.20	0.65
	Obesity	3	Lean	Overweight	11470	17390	166	1.45	0.95
			Lean	Obesity	11470	105694	1132	9.87	1.07
	Cancer	4	Healthy	Cancer	177296	100128	630	0.36	0.63
	HIV	5	Negative	Treatment	183240	343199	6984	3.81	2.03
			Negative	Non-treatment	183240	337262	3762	2.05	1.12
		6	Negative	Treatment	152069	100127	2718	1.79	2.71
			Negative	Non-treatment	152069	100127	1934	1.27	1.93
		7	Negative	Positive	77417	55596	2049	2.65	3.69
	T1D (Lean)	8	Healthy	T1D	45752	46663	1059	2.31	2.27
	T1D (Obesity)		Healthy	T1D	64261	63902	3393	5.28	5.31
	Gout	9	Healthy	Gout	34962	28916	407	1.16	1.41
	MHE	10	Healthy	MHE	19306	11175	172	0.89	1.54
			Control	MHE	24089	11175	309	1.28	2.77
	Parkinson's Disease	11	Healthy	PD	64123	173070	3630	5.66	2.10
	Schizophrenia	12	Healthy	Schizophrenia	152628	155959	4861	3.18	3.12
	Autism	13	Healthy	Autism	12245	23827	43	0.35	0.18
			Healthy	Neurotypical	12245	15043	24	0.20	0.16
	Atherosclerosis	14	Healthy	Atherosclerosis	119315	159894	4288	3.59	2.68
Oral	Atherosclerosis	15	Healthy	Atherosclerosis	1485	5253	2	0.13	0.04
	Periodontitis	16	Healthy	PB	1485	989	1	0.07	0.10
			Healthy	PnB	1485	945	1	0.07	0.11
		17	Healthy	Disease	79400	1004652	1353	1.70	0.13
			Control	Disease	735077	1004652	0	0.00	0.00
	Smoking	18	Non-smoking	Smoking	21735	15224	279	1.28	1.83
		19	Non-smoking	Smoking	21745	58222	1707	7.85	2.93
		20	Non-smoking	Smoking	5101	10841	153	3.00	1.41
Skin	Psoriasis	21	Control	Lesion	387400	485669	37098	9.58	7.64
			Normal	Lesion	488439	485669	64145	13.13	13.21
Lung	Cystic Fibrosis (CF)	22	End of Treatment	Exacerbation	1	28	0	0.00	0.00
		23	Healthy	CF	406	434	0	0.00	0.00
	HIV	24	Negative	Positive	171801	214879	19746	11.49	9.19
Vaginal	BV	25	Healthy	BV	135	6	0	0.00	0.00
Semen	Infertile (Genus level)	26	Normal	Subnormal	189	231	3	1.59	1.30
			Normal	Abnormal	189	378	6	3.17	1.59
	Infertile Species level)		Normal	Subnormal	366772	525807	26291	7.17	5.00
			Normal	Abnormal	366772	565507	24225	6.60	4.28
Mean & Standard Error			**Mean**		109477.68	164327.1	5597.37	3.05	2.27
			Standard Error		26413.27	41954.41	2086.53	0.57	0.45

Supplementary Figs S1-S3

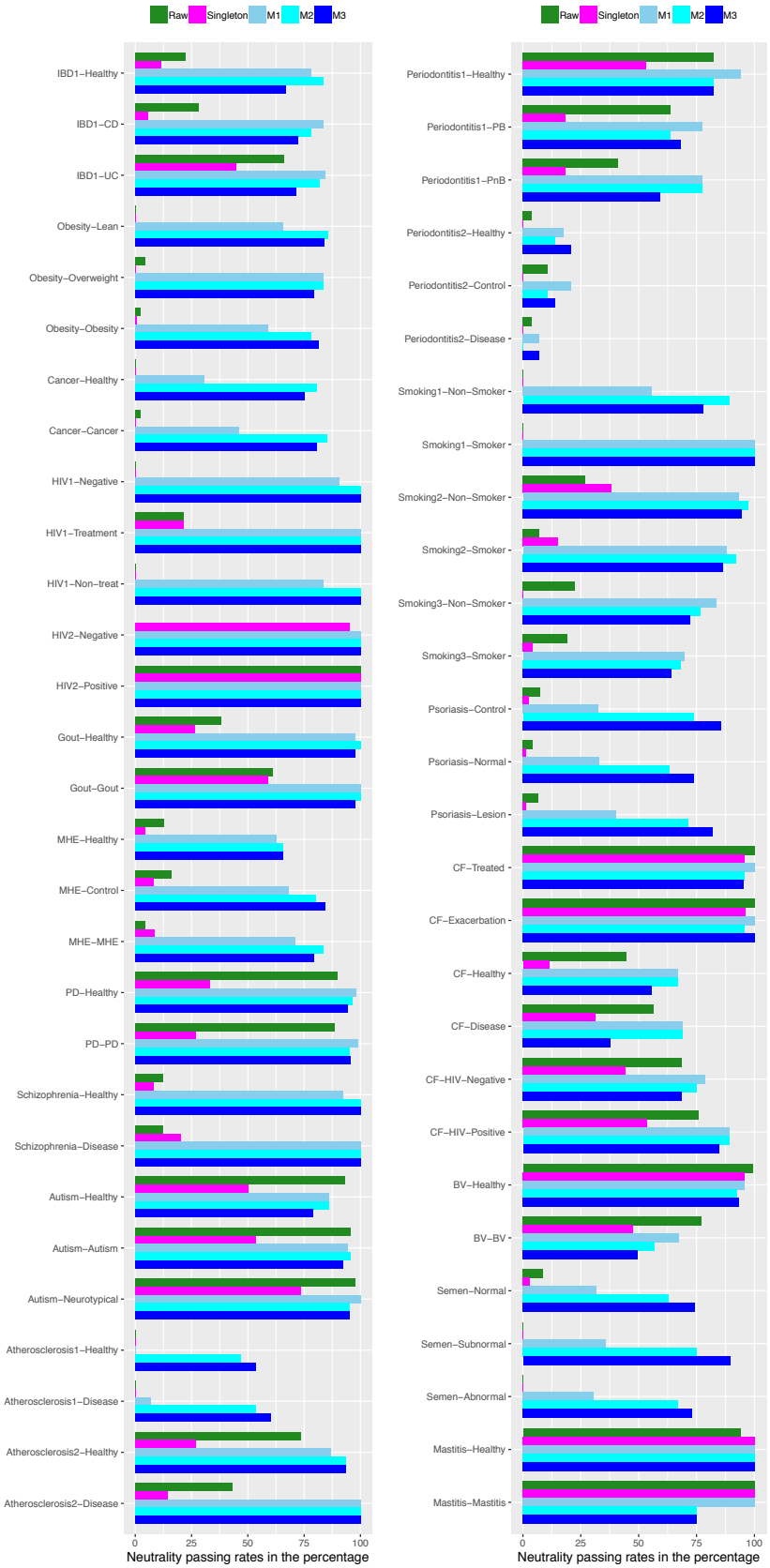

Fig S1. The neutrality passing rates for each of the 24 case studies: for each case, the test results for five data pre-processing schemes are plotted (Related to Fig 2).

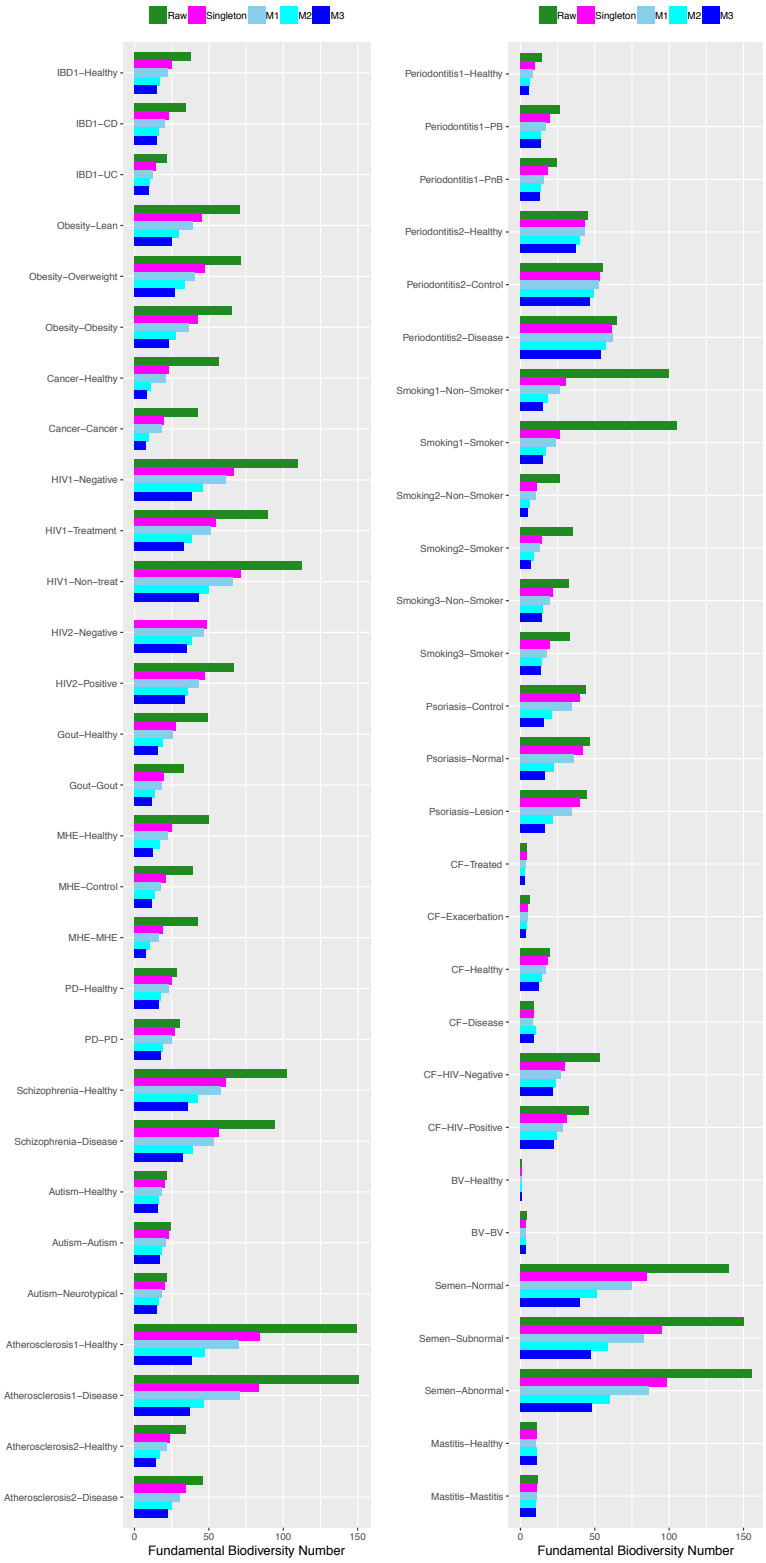

Fig S2. The fundamental diversity number (θ) for each of the 24 case studies: for each case study, the test results for five data pre-processing schemes are plotted (Related to Fig 3).

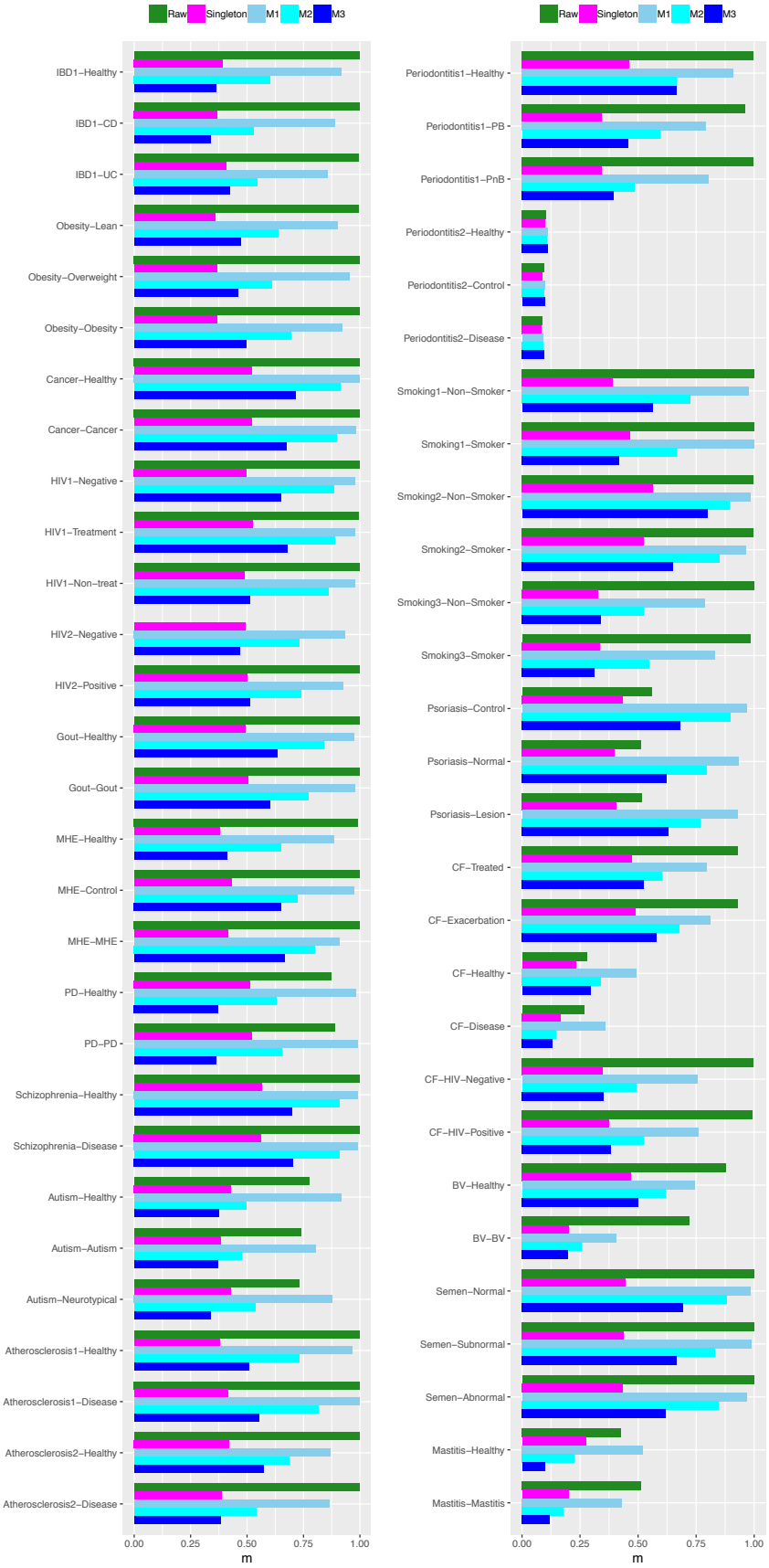

Fig S3. The immigration probability (*m*) (fundamental dispersal number) for each of the 24 case studies, for each case study, the test results for five data pre-processing schemes are plotted (Related to Fig 3).

Transparent Methods

Datasets of Human Microbiome Associated Diseases

A brief description on the metagenomic datasets of the 27 MAD (microbiome associated disease) studies is provided in Table S1 of the online supplementary information (OSI), which are the same datasets used in Ma *et al*. (2019). Among the 27 datasets, there were 3 and 2 datasets that are not suitable for the neutral-theoretic analysis and CPN/HSN analyses, respectively, and therefore were excluded from the analyses of this study. The 27 case studies cover most of the high-profile MAD including obesity, IBD, diabetes, BV (bacterial vaginosis), periodontitis, and neurodegenerative diseases. The datasets were in the abundance of 16s-rRNA reads clustered at the species level (97% similarity level), equivalent with the species abundance tables in macrobial ecology.

Hubbell's unified neutral theory of biodiversity (UNTB)

Hubbell (2001) classic neutral theory describes a local community containing J individuals, one of which, randomly chosen, dies and is replaced at every time step. Two ways the individual is replaced are: an offspring of another randomly chosen individual from the local community, occurring with probability $(1–m)$, or offspring from a randomly chosen individual from an outside pool of individuals (metacommunity) with probability m. The metacommunity is a well-mixed source pool of individual organisms; each individual has the potential to reproduce offspring migrated to the local community being studied. Neutral theory assumes that the species abundance distribution (SAD) of metacommunity is governed by a similar neutral process. The replacement individuals in the metacommunity are offspring from a randomly chosen individual in the metacommunity, or from speciation (which occurs with probability v). Obviously, in previous description of community dynamics (birth, death, migration, speciation), all individuals are assumed to be equivalent and are distinguished only by their species labels. The species labels are preserved solely for describing SAD.

The neutral theory is formulated as a sampling theory, which takes into account the effect of sampling a small proportion from the much larger natural system (sampling space). The mathematical model used to describe the neutral process is multi-nominal distribution, which can be mechanistically derived from a master equation—a differential equation that describes the community dynamics based on the neutral principles. It makes predictions about SADs, which can be statistically tested against observed species abundances data obtained from sampling actual communities. In the following, we briefly describe the parameters and statistical tests used in testing the neutrality for a sole purpose to interpret the results from our analysis.

The total number of individuals of all species in a local community (represented as $\boldsymbol{J}$) is the local community size. In our case, it is the total number of 16s-rRNA reads of all OTUs from a community sample. The UNTB assumes a finite community size.

The number of species ($\boldsymbol{S}$) in the community may change due to speciation, extinction and immigration from meta-community.

The fundamental biodiversity number (θ) is actually a measure of speciation, which can be seen from its definition, $\theta = J_M[v/(1-v)]$, where J_M is the number of individuals in the metacommunity and v is the per capita speciation rate. Therefore, the fundamental biodiversity number (θ) represents the speciation process in Vellend (2010) & Hanson et al. (2012) synthesis for community patterns.

The immigration probability (m) is a measure for *dispersal limitation*, and it is defined as $m=I/(I+J-1)$, where I is the number of immigrants that are searching for vacant spot in the local community (*i.e*., competing with the local individuals), and J is the number of individuals as mentioned previously. Dispersal limitation is a process that causes the location of an individual to be limited in some sense by the location of its parent (Rosindell *et al*. 2011). The immigration probability (m) therefore, characterizes the dispersal process in Vellend (2010) and Hanson *et al.* (2012) synthesis for community ecology.

Etienne's (2005, 2007) exact neutrality test adopts a mixture test strategy of Monte Carlo (simulation) significance test and the parametric bootstrap. The test is 'exact' since type I error can be *exactly* specified. It compares the probability (the likelihood) of the realized or observed configuration (*i.e*., SAD of actually sampled community) with the probabilities of the artificially simulated configurations based on the neutral theory model by using a Chi-squared test. If the p-value from the test is $p>0.05$, we conclude that the neutrality of the community under testing cannot be rejected, and that the observed SAD is consistent with the prediction from the neutral theory, *i.e.,* passing the neutrality test.

To apply Hubbell's (2001) UNTB for measuring the effects of drift, speciation and dispersal (Rosindell *et al*. 2011), we adopted Etienne (2005) sampling formula (distribution model) and a standard *R*-package UNTB by Hankin (2007) that implemented Etienne (2005) exact probability test procedure. An illustration of testing the neutrality of the human microbiome is provided in Li & Ma (2015).

In consideration of the potential influence of sequencing errors on the neutrality test, we removed singleton (removing the OTUs represented by a single read), minus one, two, and three reads across all OTUs, respectively. We also included the test with "raw" OTU reads, *i.e*., without removing any reads from the OTU tables computed from standard bioinformatics pipelines used in Ma *et al*. (2019). That is, a total of five tests are performed for each community sample based on five different data pre-processing schemes to enhance the robustness of our analysis.

Fisher's exact probability test is designed to compare the frequency of occurrences (observations) in a fourfold table setting when the numbers are too small to use the Chi-square test (Fisher 1922, 1954). We use Fisher's exact probability test to determine if there are pair-wise differences between the treatments (the healthy and diseased) in their passing rates of the neutrality test. We used the *R*-function (fisher.test) from the standard R-package (stats) (http://stat.ethz.ch/R-manual/R-patched/library/stats/html/fisher.test.html) to conduct the Fisher's exact test in this study. Note that the step of Fisher's exact test was performed after the FDR (false discovery rate, *i.e*., multiple testing corrections) control for the p-values from Etienne's (2005) neutrality test of Hubbell's (2001) UNTB.

Ning et al. (2019) framework for quantifying ecological stochasticity

We use a new framework recently developed by Ning *et al*. (2019) for quantifying ecological stochasticity to cross-verify the results from neutral theory modeling. The theoretical basis of their mathematical framework is that deterministic processes should drive ecological communities more similar or dissimilar than null expectation, and they formulated a sophisticated procedure to implement a null model for quantifying stochasticity. Our motivation to apply Ning *et al*. (2019) framework for estimating community stochasticity level is to cross-verify the findings from Hubbell's UNTB, especially to determine a possible ceiling limit for stochasticity given that UNTB model tended to overestimate the stochastic neutral processes.

Technically, this is possible since the normalized stochasticity ratio (NSR) of Ning *et al.* (2019) is a dimensionless metric in percentages or between *0* and *1*. The *NSR* represents for the *strength of stochasticity* in the community assembly, and should range from 0 to 100%. If the community assembly is extremely deterministic without any stochasticity, then *NSR* should be 0%; otherwise *SR* should be 100%. However, when expected stochasticity is very low, *NSR* could overestimate stochasticity (Ning *et al*. 2019).

Core/periphery network (CPN)

Given that network analysis has been widely applied to microbiome research (*e.g.,* Faust *et al*. 2012, Faust & Raes 2012, Kurtz *et al*. 2015, Ma & Ye 2017, Xiao *et al*. 2017, Röttjers & Faust 2018, Ma & Ellison 2019), what are special with CPN or HSN (*see* next sub-section)? Informally, the network core usually denotes a centrally and densely connected set of network nodes, while the network periphery refers to a sparsely connected, usually non-central set of nodes that are linked to the core (Csermely *et al*. 2013). Ecological communities/systems are typical complex systems (networks), which can be formulated as core/periphery network (CPN). More recent studies established that network cores promotes system robustness and evolvability, which can help system to adapt to large fluctuations of the environment, as well as to noise of intrinsic processes (Csermely *et al*. 2013). The core/periphery structure was found to be ubiquitous in the species dominance network of the human vaginal microbiome and the structure plays an important role in controlling the community diversity-stability relationships (Ma & Ellison 2019). For example, an extensive core (extensively large core size) allows for a more resilient network (community) but could be more difficult to control.

According to Csermely *et al*. (2013), a perfect or ideal core/periphery network consists of a fully linked core and a periphery that is fully connected to the core, but none of the periphery nodes are connected with each other. Formally, let $G=(V, E)$ be an undirected, unweighted graph with n nodes and m edges, and let $A=(a_{ij})$ is the adjacency matrix of G, where $a_{ij}=1$ if node i & node j are linked and *0* otherwise. Let δ be a vector of length n with entries of *1* or *0*, if the corresponding node belongs to the core or the periphery, respectively. Additionally, let $P=(p_{ij})$ be the adjacency matrix of the ideal or perfect core/periphery network of n nodes and m edges. The detection of core-periphery structure is an optimization problem to find the vector δ such that the objective function (ρ) achieves its maximum. With the vector δ, it is then trivial to classify nodes into either core or periphery.

$$\rho = \sum_{i,j} A_{ij} P_{ij} \qquad (1)$$

When applied to the human microbiome, the CPN structures reflect the heterogeneity or asymmetry of species (OTUs) from node perspective. In terms of the Vellend-Hanson four-processes synthesis (Vellend 2010, Hanson et al. 2012), which states that drift, speciation, dispersal, and selection are the underlying processes or mechanisms shaping community diversity patterns and driving community dynamics, the heterogeneity or asymmetry of species (nodes) revealed by CPN represents for the effects of *selection.* Therefore, at least, conceptually, the CPN analysis offers a framework to investigate the effects of selection from node perspective.

More importantly, the core/periphery (nodes or species) can be considered as critical network structures, and we postulate that they should be more sensitive to disturbances including MADs. In the present study, we mainly perform two kinds of statistical tests with CPNs. The first is to conduct shared core/periphery analysis, which assesses and interprets the effects of MADs on the possible increase or decrease of shared core/periphery between the healthy and diseased

individuals. The second is to test the differences between the healthy and diseased individuals in the properties of critical network structures.

High-salience skeleton network (HSN) and network assortativity

While core/periphery network distinguishes the different structural and functional roles between core and periphery nodes (species), the high-salience skeleton network (HSN) makes distinctions among the links (edges). Arguably the most important reason (also the advantage) why network analysis has been experiencing explosion is its capacity in reducing the *complexity* of complex systems (networks) while preserving their certain key features. In other words, network analysis can shed critical lights on complex networks by reducing complexity and leveraging the information from the key features. The HSN allows us to focus on critical paths (interactions) in complex networks (Grady *et al*. 2012, Shekhtman *et al*. 2014).

High salience skeletons or backbones reduce the number of links in the network while preserving the nodes (Grady *et al*. 2012, Shekhtman *et al*. 2014). However, reducing link complexity is particularly challenging because of the *mix* of link and node heterogeneity. In the case of human microbiome network, the heterogeneities of both nodes and links were equally significant, which is evidenced by the near universal fitting of the power law distributions to various network properties of nodes and links (Ma & Ellison 2019). Whether generic heterogeneous networks can be intrinsically segregated into qualitatively distinct clusters is still an open problem in network theory (Grady *et al*. 2012, Shekhtman *et al*. 2014). Grady *et al*. (2012) introduced the concept of link salience to deal with the problem, which measures the significance of a link and is based on an ensemble of node-specific perspectives of the network, and quantifies the extent to which a consensus among nodes exist regarding the importance of a link. The so-termed *high-salience skeletons* then constitute the backbones ("highways") of the network.

Besides identifying backbones or highways in complex networks, another advantage of high salience skeleton network is its particular suitability for exploring the response of networks under different perturbations such as intentional attack and random failure. High salience network has been particularly successful in the studies of transportation networks, where skeletons (backbones) are responsible for carrying the majority of the traffic. When encountering perturbations such as traffic jams or natural disasters, backbones again are critical for transportation rerouting and cancellations to survive disasters (Grady *et al*. 2012, Shekhtman *et al*. 2014). In this study, MADs can be treated as perturbations at ecological time scale and as selection force at evolutionary time scale, which may not be separable according to Hanson *et al*. (2012) due to the fast-track evolution of microbes.

Link salience (s) is defined based on the notion of shortest paths in weighted networks, *e.g*., the species correlation network with the *inverse* of correlation coefficients as weights (Ma & Ellison 2019). Assume a weighted network defined by weight matrix w_{ij} and a shortest path between node x and y, the indicator function can be defined as: $\sigma_{ij}(y, x) = 1$ if edge (i, j) is on the shortest path from x to y, $\sigma_{ij}(y, x) = 0$, otherwise. A shortest path tree $T(x)$ rooted at node x is described by a matrix with elements: $T_{ij}(x) = 1$, if $\sum_{y} \sigma_{ij}(y, x) > 0$, $T_{ij}(x) = 0$ otherwise. Link salience s_{ij} of edge (i, j) is computed with the following formula:

$$s_{ij} = \frac{1}{N} \sum_{x} T_{ij}(x) = \left\langle T_{ij}(x) \right\rangle_{V} \qquad (2)$$

where $\langle \bullet \rangle_{V}$ is the average across the set of root nodes x.

As explained previously, when applied to the human microbiome, the high salience skeletons (links) reflect the heterogeneity or asymmetry of species (OTUs) interactions from the link perspective. According to the four-process synthesis (Vellend 2010, Hanson *et al*. 2012), heterogeneity or asymmetry of species interactions or species interactions with their environments represents for the effects of *selection,* which can be revealed by the HSN introduced here from link perspective, cross-verified with the CPN from node perspective introduced previously.

Specifically, we will first test whether or not the *shared* high-salience skeletons between the healthy and diseased individuals are increased or decreased due to the disease effects, similar to the previously described shared core/periphery nodes analysis. Second, we will also test the properties of high-salience skeleton networks, which are critical network structures from link perspective, similar to the testing of properties of core/periphery network from node perspective. Furthermore, integrated with previous CPN analysis, the combined *shared analyses* of the core/periphery/skeleton structures, as well as testing the differences in the key properties of the core/periphery/skeleton networks, are expected to reveal the differences between the healthy and diseased individuals and identify potentially invaluable diagnosis and risk prediction indicators for MADs.

As to the CPN/HSN properties, we mention one particularly important network property, the *assortativity* coefficient. The property is important because it is highly relevant to network resilience, the loss of which is known as *dysbiosis* in the literature of human MADs. In assortative networks, the nodes with many connections tend to be connected to other nodes that also possess many connections (Newman 2002, 2003). The assortativity coefficient (r) is a concept for measuring the homophyly level of the graph (network). The homophyly refers to resemblance due to common ancestry based on values (such as degree) assigned to vertices. A higher assortativity coefficient (r) indicates that connected vertices tend to have more similar assigned values. For an undirected network, the assortativity based on vertex degree is defined as

$$r = \frac{1}{\sigma_q^2} \sum_{jk} jk(e_{jk} - q_j q_k), \qquad (3)$$

where e_{jk} is the fraction of edges connecting vertices of type j and k (*i.e.*, vertices that have degrees j and k), σ_q is the standard deviation of the distribution q_k, and q_k is equal to: $q_k = \sum_j e_{jk}$.

From eqn. (3), one can find that the r is equal to the Pearson's correlation coefficient of degree between pairs of linked nodes. The value of r lies in the range $-1 \le r \le 1$, with $r = 1$ indicating perfect assortativity and $r = -1$ indicating perfect disassortativity. When the value of r lies around 0, it is null or neural result, meaning that it is not statistically different from zero. Assortative networks are more resilient against removal of their vertices, especially highest-degree vertices, than disassortative or neutral ($r \approx 0$) networks. The above definition for assortativity index is based on general complex network such as species correlation network (SCN). In this study, we compute the assortativity (r_{HSS}) for the high-salience skeleton networks based on the degree of high-salience skeletons (HSS) of vertices. Different from SCNs, the high-salience skeleton networks are typically disassortative (Grady 2012).

Computational procedures/algorithms for CPN & HSN analyses

As introduced previously, two types of statistical tests, *i.e*., the shared core/periphery/skeleton analyses between the healthy (H) and diseased (D) treatments (*i.e.,* test the change of shared critical network structures between the H & D treatments), as well as test the differences in key properties of the critical network structures between the H & D treatments. Both types of tests depend on the construction of CPN/HSN, which are built from species correlation networks

(SCN). The samples from both H & D treatments are randomly mixed before building SCN networks. From the randomly mixed (permutated) samples, 1000 pairs (pair=H *vs*. D) of permutated networks are constructed, and the *p*-values of the permutation tests for determining the significances of the differences (*i.e*., disease effects) are computed from comparing the 1000 pairs of networks. Spearman's correlation coefficients with FDR (false discovery rate) control with *p*-value=0.001 are first used to construct SCNs. From the SCNs, the algorithms outlined in previous sections are applied to reconstruct the corresponding CPN and HSN networks. The computation procedures for building SCN can be found in Junker & Schreiber (2008) and Ma & Ye (2017), and those for building CPN/HSN can be found in Ma & Ellison (2019). Further information on the algorithms for performing the two types of statistical tests outline above is briefly introduced below:

General outline for generating permutated networks for performing shared core/periphery/skeleton analyses and for testing the network properties

(1) Building 1000 sets (pairs) of healthy (*H*) & diseased (*D*) networks by randomly reshuffling the *H* and *D* samples for 1000 times using Spearman's rank correlation coefficients with FDR (false discovery control).

A. For Shared Core/Periphery (Nodes) Analysis
(2A) Detect core/periphery nodes from the *H* & *D* networks built in (*1*), construct core/periphery networks (CPN), and compute shared core/periphery nodes between the *H* & *D* samples.
(3A) Perform permutation test with the 1000 permutated networks and test the change of shared core/periphery nodes between the *H* & *D* samples, based on the results from (*2A*).

B. For Shared Skeleton (Edge or Link) Analysis
(2B) Detect high salience skeletons from the H & D networks built in (*1*), construct high salience skeleton networks (HSN), and compute the shared skeletons between the H & D samples.
(3B) Perform permutation test with the 1000 permutated networks and test the change of shared skeletons between the H & D samples, based on the results from (*2B*).

C. Significance Test for Core/Periphery and Skeleton Networks Properties
Using the results from A & B to test the difference in the network properties by performing the standard permutation tests.

Test Algorithms

Algorithm for Shared Core/Periphery Analysis based on Permutated Networks
Step [1]: Assuming there are *a* samples from the healthy (*H*) treatment, and *b* samples from the diseased (*D*) treatment, compute the pair-wise Spearman's correlation coefficients for the *a* samples in the H and the *b* samples in the *D* treatment, respectively. Use FDR control with *p*=0.001 to filter out insignificant correlations, and obtain the final correlation relationships for building the species correlation networks for the *H* and *D* treatments. From the two species correlation networks, detect the core/periphery nodes for each network, and compute the shared core/periphery nodes between the *H* & *D* networks, respectively. This step is no difference from regular core/periphery network analysis, until computing the shared core/periphery nodes.

The shard core/periphery nodes obtained from this *Step* [1] is the actually or *observed* shared core/periphery nodes.

Step [2]: Pool together all samples from *H* & *D* and perform random permutation of the combined (*a*+*b*) samples. Treat the first *a* samples as the permutated healthy treatment and the leftover *b* samples as the permutated diseased treatment.

According to the algorithm in *Step* [1], compute the shared core/periphery nodes for this specific pair of permutated *H* & *D* treatments.

Step [3]: Repeat *Step* [2] for 1000 times, and obtain 1000 sets of shared core/periphery nodes.

Step [4]: Compute a pseudo-*p* value. Using the number of shared core nodes as example, assume the shared core nodes from **Step [1]** is *N*, the numbers of shared core nodes from the 1000 random permutations in *Step* [2-3] are $N_1, N_{2,...,} N_{i,...}N_{1000}$, the pseudo-*p* value is the proportion of the permutations with $N_i>N$ among 1000 times of permutations. That is, assuming *n* is the number of times satisfying $N_i>N$ in 1000 permutations, $p=n/1000$.

Algorithm for Shared Skeleton Analysis Based on Permutated Networks
Step [1]: Assuming there are *a* samples from the healthy (*H*) treatment, and *b* samples from the diseased (*D*) treatment, compute the pair-wise Spearman's correlation coefficients for the *a* samples in the H and the *b* samples in the *D* treatment, respectively. Use FDR control with $p=0.001$ to filter out insignificant correlations, and obtain the final correlation relationships for building the species correlation networks for the *H* and *D* treatments. From the two species correlation networks, compute the HSS (high salient skeleton) value for each edge, and compute the shared skeletons between the *H* & *D* networks with HSS>0, and HSS≥0.2, respectively. This step is no difference from regular skeleton network analysis, until computing the shared skeletons.

The shard skeletons obtained from this *Step* [1] is the actually or *observed* shared skeletons.

Step [2]: Pool together all samples from *H* & *D* and perform random permutation of the combined (*a*+*b*) samples. Treat the first *a* samples as the permutated healthy treatment and the leftover *b* samples as the permutated diseased treatment.

According to the algorithm in *Step* [1], compute the shared skeletons for this specific pair of permutated *H* & *D* treatments.

Step [3]: Repeat *Step* [2] for 1000 times, and obtain 1000 sets of shared skeletons.

Step [4]: Compute a pseudo-*p* value. Assume the shared skeletons from **Step [1]** is *N*, the numbers of shared skeletons from the 1000 random permutations in Step [2-3] are $N_1, N_{2,...,} N_{i,...}N_{1000}$, the pseudo-*p* value is the proportion of the permutations with $N_i>N$ among 1000 times of permutations. That is, assuming *n* is the number of times satisfying $N_i>N$ in 1000 permutations, $p=n/1000$.

References (Cited in "Transparent Methods")

Abusleme L, Dupuy AK, Dutzan N, et al. (2013) The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *The ISME Journal*, 7(5):1016-1025.

Alekseyenko AV, Perezperez GI, Souza AD, et al. (2013) Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome*,1,1(2013-12-23), 1(1):31-31.

Blainey PC, Milla CE, Cornfield DN, et al. (2012) Quantitative analysis of the human airway microbial ecology reveals a pervasive signature for cystic fibrosis. *Science Translational Medicine*, 4(153):153ra130.

Charlson ES, Chen J, Custersallen R, et al. (2010). Disordered microbial communities in the upper respiratory tract of cigarette smokers. *Plos One*, 5(12):e15216.

Csermely P, London A, Wu LY, Uzzi B. Structure and dynamics of core/periphery networks. *Journal of Complex Networks.* 2013;1:93-123.

Dilip V. Jeste (UC San Diego) https://profiles.ucsd.edu/dilip.jeste

Etienne RS (2005) A new sampling formula for neutral biodiversity. *Ecology Letters* 8(3):253-260.

Etienne RS. (2007) A neutral sampling formula for multiple samples and an 'exact' test of neutrality. *Ecology Letters* 10(7): 608-618.

Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J et al. (2012) Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol*. 2012;8(7):e1002606.

Faust K, Raes J. (2012) Microbial interactions: from networks to models. *Nat Rev Microbiol*. 10:538–550.

Fisher RA (1922) On the interpretation of χ^2 from contingency tables, and the calculation of p-value. *Journal of the Royal Statistical Society*. Vol. 85(1): 87–94.

Fisher RA (1954) Statistical Methods for Research Workers. *Oliver and Boyd*. ISBN 0-05-002170-2.

Fodor AA, Klem ER, Gilpin DF, et al. (2012) The Adult Cystic Fibrosis Airway Microbiota Is Stable over Time and Infection Type, and Highly Resilient to Antibiotic Treatment of Exacerbations. *Plos One*, 7(9):e45001.

Grady D, Thiemann C, Brockmann D. (2012) Robust classification of salient links in complex networks. *Nature Communications.* 2012;DOI: 10.1038/ncomms1847

Griffen AL, Beall CJ, Campbell JH, et al. (2012) Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *The ISME Journal*, 6(6):1176.13

Guo Z, Zhang J, Wang Z, et al. (2016) Intestinal Microbiota Distinguish Gout Patients from Healthy Humans. Sci Rep, 6:20602.

Halfvarson J, Brislawn CJ, Lamendella R, et al. (2017) Dynamics of the human gut microbiome in Inflammatory Bowel Disease. *Nature Microbiology*, 2:17004.

Hankin RKS (2007) Introducing UNTB, an R Package For Simulating Ecological Drift Under the Unified Neutral Theory of Biodiversity. *Journal of Statistical Software*. 22(12):15pp. (https://cran.r project.org/web/packages/untb/index.html)

Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol*. Vol.10:497–506

Hill-Burns EM, Debelius JW, Morton JT et al. (2017) Parkinson's Disease and PD Medications Have Distinct Signatures of the Gut Microbiome. *Movement Disorders Official Journal of the Movement Disorder Society*, 32(5):739.

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. *Princeton University Press*.

Junker BH, Schreiber F. (2008) *Analysis of biological networks*. Wiley. 2008.

Kang DW, Adams JB, Gregory AC, et al. (2017) Microbiota Transfer Therapy alters gut ecosystem and improves gastrointestinal and autism symptoms: an open-label study. *Microbiome*, 5(1):10

Kim, Jane (https://clinicaltrials.gov/ct2/show/NCT02938806)

Koren O, Klaenhammer TR. (2011) Human oral, gut, and plaque microbiota in patients with atherosclerosis. *PNAS*, 108(Suppl 1):4592-4598.

Kurtz ZD, Müller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA. (2015) Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLOS Comp. Biol*. 2015. https://doi.org/10.1371/journal.pcbi.1004226

Lazarevic V, Whiteson K, Hernandez D, et al. (2010) Study of inter- and intra-individual variations in the salivary microbiota. *BMC Genomics*, 11(1):1-11.

Li LW, Ma ZS (2016) Testing the Neutral Theory of Biodiversity with Human Microbiome Datasets. *Scientific Reports*. 2016;6, Article No. 31448

Lozupone C, Cota-Gomez A, Palmer B E, et al. (2013) Widespread colonization of the lung by Tropheryma whipplei in HIV infection. *American Journal of Respiratory & Critical Care Medicin*e, 187(10):1110-1117.

Ma ZS, Ye D. (2017) Trios—promising *in silico* biomarkers for differentiating the effect of disease on the human microbiome network. *Sci Rep* 2017;**7**:13259.

Ma ZS, Ellison AM (2019) Dominance network analysis provides a new framework for studying the diversity-stability relationship. *Ecological Monographs*. https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/ecm.1358

Ma ZS, Li LW, Gotelli NJ (2019) Diversity-disease relationships and shared species analyses for human microbiome-associated diseases, *The ISME Journal*. 2019;DOI: 10.1038/s41396-019-0395-y

Mchardy IH, Li X, Tong M, et al. (2013) HIV Infection is associated wth compositional and functional shifts in the rectal mucosal microbiota. *Microbiome*, 1(1):26.

Neff CP, Krueger O, Xiong K, et al. (2018) Fecal Microbiota Composition Drives Immune Activation in HIV-infected Individuals. *Ebiomedicine*, 30

Newman, MEJ (2002) Assortative mixing in networks. *Physical Review Letters*, 89(20), 208701.

Newman MEJ (2003) Mixing patterns in networks. *Phys. Rev. E* 67, 026126

Ning D, Deng Y, Tiedje JM, Zhou J (2019) A general framework for quantitatively assessing ecological stochasticity. *PNAS*, 116 (34): 16892-16898.

Papa E, Docktor M, Smillie C, et al. (2012) Non-Invasive Mapping of the Gastrointestinal Microbiota Identifies Children with Inflammatory Bowel Disease. *Plos One*, 7(6):e39242.

Rosindell J, Hubbell SP, Etienne RS. (2011) The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten. *Trends in Ecology & Evolution.* Vol. **26**:340–348.

Röttjers L, Faust K (2018) From hairballs to hypotheses–biological insights from microbial networks. *FEMS Microbiology Reviews*, 42(6):761–780. https://doi.org/10.1093/femsre/fuy030

Shekhtman LM, Bagrow JP, Brockmann D. (2014) Robustness of skeletons and salient features in networks. *Journal of Complex Networks*. doi:10.1093/comnet/cnt019

Srinivasan S, Hoffman NG, Morgan MT, et al. (2012) Bacterial Communities in Women with

Bacterial Vaginosis: High Resolution Phylogenetic Analyses Reveal Relationships of Microbiota to Clinical Criteria. *Plos One*, 7(6):e37818.

Turnbaugh PJ, Hamady M, Yatsunenko T, et al. (2009) A core gut microbiome in obese and lean twins. *Nature*, 457(7228):480.

Urbaniak C, Mcmillan A, Angelini M, et al. (2014) Effect of chemotherapy on the microbiota and metabolome of human milk, a case report. *Microbiome*, 2(1):24.

Vellend M (2010) Conceptual synthesis in community ecology. *Q. Rev. Biol*. 2010;85:183–206

Wang T, Cai G, Qiu Y, et al. (2012) Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. ,*The ISME Journal* 6(2):320-329.

Weng SL, Chiu CM, Lin FM, et al. (2014) Bacterial Communities in Semen from Men of Infertile Couples: Metagenomic Sequencing Reveals Relationships of Seminal Microbiota to Semen Quality. *Plos One*, 9(10):e110152.

Xiao Y, Angulo MT, Friedman J, Waldor MK, Weiss ST, Liu YY. (2017) Mapping the ecological networks of microbial communities. *Nature Communications*. Vol. **8**, Article number: 2042

Zhang Z, Zhai H, Geng J, et al. (2013) Large-scale survey of gut microbiota associated with MHE Via 16S rRNA-based pyrosequencing. *American Journal of Gastroenterology*, 108(10):1601-1611.